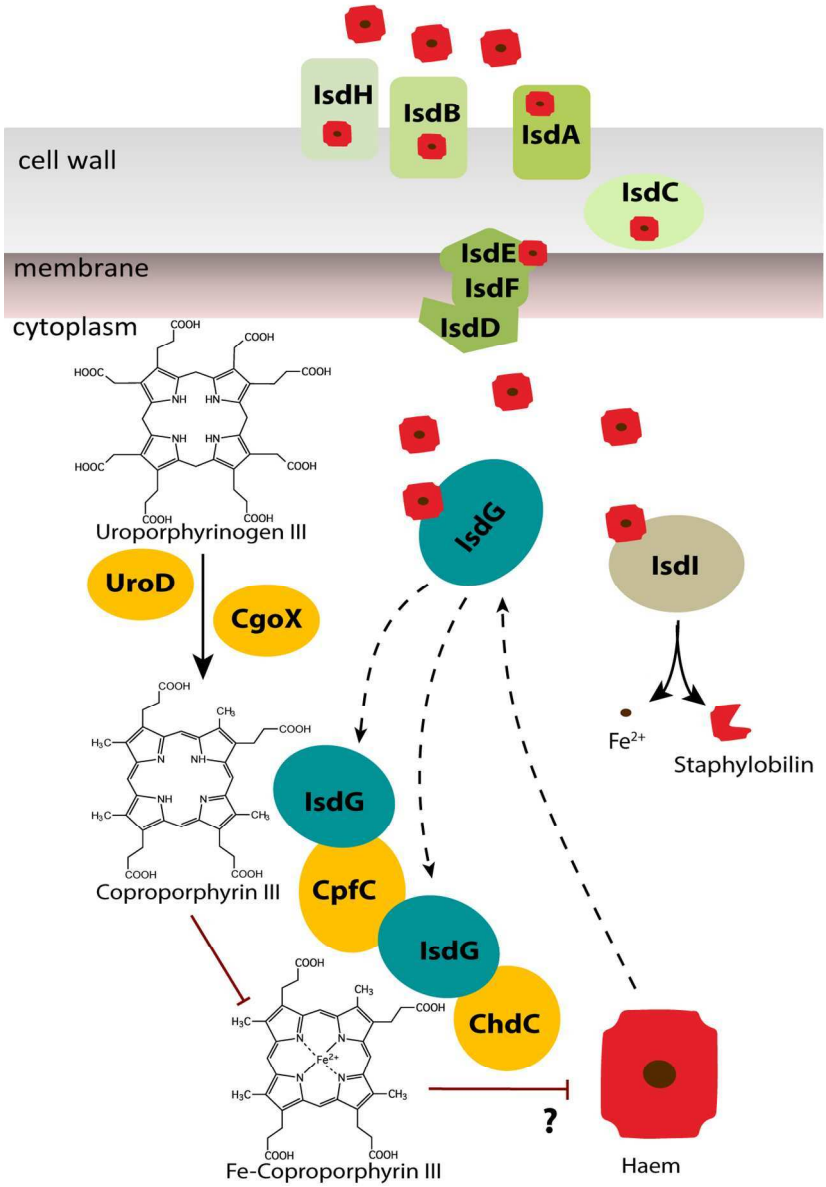
Staphylococcus aureus haem biosynthesis and acquisition pathways are linked by haem monooxygenase IsdG

Journal:	*Molecular Microbiology*
Manuscript ID	MMI-2018-16976.R2
Manuscript Type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Videira, Marco; Instituto Tecnologia Química e Biológica, Biologica Chemistry Lobo, Susana; Universidade Nova de Lisboa, Instituto de Tecnologia Química e Biológica António Xavier Silva, Liliana; Instituto Tecnologia Química e Biológica, Biologica Chemistry Palmer, David; University of Kent, Biological Sciences Warren, Martin; University of Kent, Biosciences; Prieto, Manuel; Universidade de Lisboa Instituto Superior Tecnico, Chemistry Coutinho, Ana ; Universidade de Lisboa Instituto Superior Tecnico, Chemistry Sousa, Filipa; University Wien, Biology Fernandes, Fabio; Universidade de Lisboa Instituto Superior Tecnico, Chemistry Saraiva, Ligia; Instituto Tecnologia Química e Biológica, Biologica Chemistry;
Key Words:	haem biosynthesis, ferrochelatase, S. aureus, haem uptake, haem monooxygenase

Abbreviated Summary

Haem is an essential cofactor that in pathogens is obtained by means of an internal haem biosynthesis pathway and/or from the host by a capture system. Here it is shown for the first time that pathogens possessing the two systems communicate through a haem monooxygenase enzyme of the IsdG-family type. The importance of this crosstalk lies in the need for pathogens to reconcile their high haem requirements with the paradox that free haem is toxic.



111x163mm (300 x 300 DPI)

Staphylococcus aureus haem biosynthesis and acquisition pathways are linked through haem monooxygenase IsdG

Marco A.M. Videira[1], Susana A.L. Lobo[1,2], Liliana S.O. Silva[1], David J. Palmer[3], Martin J. Warren[3], Manuel Prieto[4], Ana Coutinho[4,5], Filipa L. Sousa[6], Fábio Fernandes[4,7], and Lígia M. Saraiva[1]*

[1] Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, Avenida da República , 2780-157 Oeiras, Portugal.

[2] iBET, Instituto de Biologia Experimental e Tecnológica, Oeiras, Portugal. Avenida da República (EAN), 2780-157 Oeiras, Portugal (present address).

[3] School of Biosciences, University of Kent, Giles Lane, Canterbury, Kent CT2 7NJ, United Kingdom.

[4] Centro de Química-Física Molecular and Institute of Nanoscience and Nanotechnology, Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal.

[5] Departamento de Química e Bioquímica, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal.

[6] Department of Ecogenomics and Systems Biology, University of Vienna, 1090 Vienna, Austria.

[7] Research Unit on Applied Molecular Biosciences–Rede de Química e Tecnologia (UCIBIO-REQUIMTE), Departamento de Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal.

Corresponding author:

Instituto de Tecnologia Química e Biológica António Xavier, UNL

Av. da República (EAN)

2780-157 Oeiras

Portugal

Tel: +351 21 4469328

Fax: +351 21 4433 644

email: lst@itqb.unl.pt

Abstract

Haem is an essential cofactor in central metabolic pathways in the vast majority of living systems. Prokaryotes acquire haem via haem biosynthesis pathways, and some also utilize haem uptake systems, yet it remains unclear how they balance haem requirements with the paradox that free haem is toxic. Here, using the model pathogen *Staphylococcus aureus*, we report that IsdG, one of two haem oxygenase enzymes in the haem uptake system, inhibits the formation of haem via the internal haem biosynthesis route. More specifically, we show that IsdG decreases the activity of ferrochelatase and that the two proteins interact both *in vitro* and *in vivo*. Further, a bioinformatics analysis reveals that a significant number of haem biosynthesis pathway containing organisms possess an IsdG-homologue and that those with both biosynthesis and uptake systems have at least two haem oxygenases. We conclude that IsdG-like proteins control intracellular haem levels by coupling the two pathways. IsdG is thus a novel target for the treatment of *S. aureus* infections.

Introduction

In all forms of life, the ability to perform redox reactions is crucial, often requiring compounds with metal centres and cofactors that are evolutionarily ancient. Haem is an iron-based redox centre that is widely distributed in living organisms and which permits an extensive range of proteins and enzymes to function in essential cellular processes such as aerobic and anaerobic respiration, detoxification, microRNA processing and regulation of gene expression (Choby and Skaar, 2016). Haem is obtained by either *de novo* synthesis (via haem biosynthesis pathways) or by uptake from the environment. Several pathogens extract haem from their host's haemoglobin, which upon degradation serves as a source of iron to satisfy the nutritional needs of the pathogen (Anzaldi and Skaar, 2010; Choby and Skaar, 2016). Furthermore, a large number of microorganisms combine pathways for the endogenous production of haem with complex haem uptake systems.

Prokaryotes utilise one of three distinct haem biosynthesis pathways, all of which have common early steps that lead to the intermediate uroporphyrinogen III (Fig. 1). Thereafter, the three distinct routes are referred to as the protoporphyrin, sirohaem and coproporphyrin branches, reflecting intermediates that are unique to each pathway, and the genes involved were recently renamed (Dailey *et al.*, 2017). The protoporphyrin dependent (PPD) branch, previously known as classic pathway, operates in many Gram-negative bacteria. This pathway involves the decarboxylation of uroporphyrinogen III to coproporphyrinogen III by UroD (formerly HemE), followed by decarboxylation to protoporphyrinogen IX by CgdH/CgdC (formerly HemN/HemF) and then oxidation to protoporphyrin IX by PgoX/PgdH1 (formerly HemY/HemG) prior to metal insertion by ferrochelatase PpfC (formerly HemH) to give protohaem (Heinemann *et al.*, 2008; Dailey *et al.*, 2017). The sirohaem dependent branch (SHD), originally named the alternative haem biosynthesis pathway Ahb, is active in sulphate-reducing proteobacteria and archaea, transforms uroporphyrinogen III to protohaem via sirohaem in an oxygen-independent process requiring four-enzymatic steps involving the enzymes AhbA-D (Bali *et al.*, 2011; Lobo *et al.*, 2012). The most recently discovered branch, namely the coproporphyrin dependent (CPD) pathway is active in several Gram-positive organisms including *S. aureus* (Lobo *et al.*, 2015). This pathway involves the conversion of uroporphyrinogen III to coproporphyrinogen III by UroD (formerly HemE) and subsequent oxidation to coproporphyrin III by CgoX (formerly HemY). Insertion of iron into coproporphyrin III to make coprohaem is mediated by CpfC (formerly HemH), and the final step involves the decarboxylation of coprohaem to protohaem by ChdC (formerly HemQ) (Dailey *et al.*, 2015; Lobo *et al.*, 2015). Although these ferrochelatases catalyse the insertion of iron in a porphyrin ring, the substrate that they act upon is different, being protoporphyrin IX for PPD, coproporphyrin III for CPD and sirohydrochlorin for SHD pathways.

The mechanisms of haem biosynthesis in prokaryotes vary widely (Dailey *et al.*, 2017) and control over the pathway has been reported to be regulated by iron, oxygen, reactive oxygen species and by its final product haem (Choby and Skaar, 2016; Dailey *et al.*, 2017). In the case of the latter, the binding of haem to GtrR (formerly HemA), the first enzyme of the pathway, is thought to control the system via a

feedback inhibition mechanism (McNicholas *et al.*, 1997; Wang *et al.*, 1997; Choby and Skaar, 2016).The most common haem acquisition system in Gram-positive pathogens is the iron-regulated surface determinant (Isd). *S. aureus* also contains an additional haem transport system (Hts) that, like the Isd proteins, was first described in this bacterium (Mazmanian *et al.*, 2003; Skaar EP, Humayun M, Bae T, DeBord KL, 2004). The Isd system includes the cell wall-anchored proteins IsdA, IsdB, IsdC and IsdH, as well as the haem-binding protein and permease components IsdD, IsdE, and IsdF, and the transmembrane ABC transporter HtsABC (Mazmanian *et al.*, 2003; Skaar EP, Humayun M, Bae T, DeBord KL, 2004). Once haem is transported into the cytoplasm, it is either degraded by the haem monooxygenases IsdG and IsdI to release iron (Skaar *et al.*, 2004), or becomes bound to membrane-associated proteins to act as a cofactor in, for example, electron transport (Thöny-Meyer, 1997). *S. aureus* also possesses the haem regulated transporter (HrtAB), an efflux pump that protects cells from haem toxicity (Stauff *et al.*, 2008).

S. aureus is clinically significant as it is responsible for a large number of human infections including respiratory, urogenital, skin burn-associated and systemic infections (Tarai *et al.*, 2013). It is also a widely studied Gram-positive pathogen and the paradigm for the CPD pathway operating alongside a haem acquisition system. Genes encoding enzymes of the CPD pathway are organized into two operons in *S. aureus*, *gtrR-hemX-hmbS-uroS-pbgS* and *uroD-cpfC-cgoX*, whereas *chdC* is isolated elsewhere in the genome (Lobo *et al.*, 2015). The *S. aureus* haem acquisition system is formed by single units *isdA, isdB, isdH, isdI* and the operon *isdCDEF-srtB-isdG* (Mazmanian *et al.*, 2003; Skaar and Schneewind, 2004). Like many other organisms, *S. aureus* encodes two haem monooxygenases, IsdG and IsdI, which share approximately 70% amino acid identity and which bind and degrade haem to produce iron and staphylobilin (Skaar *et al.*, 2004; Reniere *et al.*, 2010).

In this work, we investigated how *S. aureus* controls intracellular haem levels. We show that the haem monooxygenase, IsdG, provides a link between the haem biosynthesis and uptake pathways in *S. aureus* and that this protein prevents excessive formation of toxic haem inside bacterial cells.

Results

The haem monooxygenase IsdG interferes with haem formed via the haem biosynthesis pathway

To analyse the role of the two haem monooxygenases of *S. aureus*, the *S. aureus* CPD pathway was reconstituted in a haemin auxotrophic strain of the Gram-negative bacterium *E. coli* that is deficient in protoporphyrin ferrochelatase Δ*ppfC* (formerly Δ*visA* or Δ*hemH*). As *E. coli* does not take up exogenous haem through an Isd system, this host represents a good model to study the role of IsdG/I enzymes without the interference of an endogenous haem acquisition system, thus avoiding the use of *S. aureus* strains containing multiple mutations. The genes encoding the last three enzymes of the *S. aureus* CPD pathway, *cgoX-cpfC-chdC*, were cloned by the Link and Lock method (McGoldrick *et al.*, 2005), expressed in the presence of either IsdG or IsdI in *E. coli* Δ*ppfC*, and growth of the complemented strain was monitored. The results showed that expression of CgoX-CpfC-ChdC alone abolished the haem auxotrophy of the strain, which grows similar to an *E. coli* wild type strain. However, in the presence of either IsdG or IsdI, the growth of the strain was impaired (Fig. 2A).

In order to understand the growth behaviour of the complemented strains, the tetrapyrrole products formed during the expression of *S. aureus cgoX-cpfC-chdC* genes in the presence of IsdI or IsdG were analysed by HPLC-MS. As expected, cells expressing *S. aureus* CgoX-CpfC-ChdC exhibited a peak with a mass-to-charge ratio (*m/z*) of 616, confirming their ability to produce haem. However, this peak significantly decreased when IsdG or IsdI were also expressed (Fig. 2B), with cells expressing *S. aureus* CgoX-CpfC-ChdC-IsdI containing approximately 3 times less haem, and cells expressing CgoX-CpfC-ChdC together with IsdG showing no detectable peak corresponding to haem. These results suggest that the haem monooxygenases interfere with the haem formation *in vivo*.

The lack of growth in the presence of IsdG/I enzymes indicated that no haem was being formed, which could result from degradation and/or inhibition of its formation. To test these possibilities, we analysed the growth of *E. coli* Δ*ppfC*, complemented with its own ferrochelatase PpfC and in the presence of IsdI or IsdG. Under these conditions where the haem is formed through the PPD pathway, it was observed that only IsdI impaired the growth while IsdG had no effect (Fig. 2C).

Hence, we conclude that IsdI always acts as a haem monoxygenase as it degrades haem independently of whether it is formed by the CPD or PPD pathways while IsdG impairs the growth of the strain using the CPD pathway but not of the strain using the PPD pathway suggesting that IsdG is not acting as a haem degrading enzyme.

IsdG decreases the ferrochelatase activity of CpfC

To understand how the *S. aureus* IsdG and IsdI constrain the CPD pathway, we tested whether they interfere with the activity of the *S. aureus* ChdC enzyme that catalyses the last step of the CPD branch promoting the conversion of iron-coproporphyrin III into haem in the presence of its electron acceptor

(Lobo *et al.*, 2015). Thus, *S. aureus* ChdC was incubated with iron-coproporphyrin III in the presence of either IsdG or IsdI and hydrogen peroxide. However, control experiments performed in the absence of ChdC showed that hydrogen peroxide activates the inherent peroxidase activity in IsdG and IsdI, leading to degradation of the iron-coproporphyrin III substrate (Fig. S1). We therefore shifted our attention to the penultimate enzyme of the CPD branch, the ferrochelatase CpfC, which plays an important role in the survival of *S. aureus* (Fig. S2), as also shown for the GtrR and PbgS enzymes involved in the first steps of the CPD pathway (Hammer *et al.*, 2013; Hammer *et al.*, 2014).

In *S. aureus* and other Gram-positive operating via the CPD pathway, the ferrochelatase incorporates iron into coproporphyrin III to form iron-coproporphyrin III, which is then sequentially decarboxylated by ChdC to yield protohaem. We observed that the cell lysate of *S. aureus cpfC* mutant contained very low levels of cellular catalase activity, which was restored to wild type level upon addition of external haemin (Fig. S3). It was previously reported that the total cellular catalase activity was lower in *S. aureus pbgS* and *chdC* defective strains (Mayfield *et al.*, 2013). We next tested whether the haem monooxygenases impaired ferrochelatase activity. *S. aureus* CpfC, IsdG and IsdI were recombinantly produced and purified, yielding stable proteins that had the expected molecular masses of approximately 35 and 12.5 KDa, respectively (data not shown). *S. aureus* ferrochelatase activity was determined by following the depletion of coproporphyrin III, which yielded a specific activity of 234 ± 14 $nmol.min^{-1}mg^{-1}$. While addition of IsdI to the reaction mixture caused no significant alteration (Fig. 3A), the presence of IsdG led to a ~50% decrease in ferrochelatase activity. Furthermore, the activity of the enzyme decreased with increasing concentration of IsdG (Fig 3B).

These results show that IsdG, but not IsdI, attenuates the activity of the *S. aureus* CpfC. Therefore, we next determined the haem cellular content in cells forming haem by the CPD pathway in the presence and absence of IsdG. For this purpose, *S. aureus* wild type and Δ*isdG* mutant cells were grown in the presence of aminolaevulinic acid, the first stable precursor of the haem biosynthesis pathway. Figure 3C shows that the IsdG inactivated strain has significantly higher haem content. Therefore, we conclude that IsdG impairs CpfC lowering the amount of haem formed by the CPD pathway.

Protein-protein interaction studies between IsdG and CpfC

We reasoned that the inhibition of the *S. aureus* CpfC activity by IsdG is most likely a result from a direct interaction between the two proteins. Three type of experiments were done; pull down assays, Förster resonance energy transfer (FLIM-FRET) and fluorescence polarization binding experiments.

For the pull down assays, purified CpfC-His tagged protein bound to a Ni-Sepharose column was incubated with cells extracts expressing IsdG or IsdI. The bound proteins were eluted with imidazole containing buffer and separated electrophoretically. As visualised by SDS-PAGE, IsdG is pulled down

212 together with CpfC, while IsdI does not bind to the column-loaded CpfC (Fig. 4A). Hence, the results
213 show that there is a specific interaction between CpfC and IsdG.
214
215 To further study this interaction *in vivo* fluorescence lifetime imaging based Förster resonance energy
216 transfer (FLIM-FRET) was used. Additionally, interactions among the various CPD pathway protein
217 components and *S. aureus* haem monooxygenases were also evaluated. In order to do this, we
218 constructed *S. aureus* CpfC and ChdC proteins fused to a GFP donor molecule and IsdG, IsdI, ChdC and
219 CpfC fused to a mCherry acceptor molecule (Table S1).
220 The interaction between IsdG and CpfC was analysed in bacterial cells expressing *S. aureus* CpfC-GFP and
221 IsdG-mCherry. CpfC-GFP exhibited a mono-exponential fluorescence decay with a lifetime $\tau_{CpfC\text{-}EGFP}$ = 2.2
222 ± 0.1 ns (Fig S4). In the presence of IsdG-mCherry, the fluorescence decay of CpfC-GFP was no longer
223 accurately described by a single exponential and a shorter component had to be included in the fit,
224 indicating the presence of FRET and thus an interaction between the two proteins (Fig. 4 and Fig S4). The
225 FRET efficiencies (E) were dependent on the concentration of IsdG, according to the plots of *E* against
226 mCherry/GFP fluorescence intensity ratios (Fig. 4), obtained from the confocal fluorescence
227 measurements of each construct (see Methods).
228 Although some degree of interaction between CpfC-GFP and IsdI-mCherry was also detected (Fig. 4B
229 and 4D), the FRET efficiencies determined for this protein pair were much lower than those measured
230 for IsdG and CpfC (Fig. 4C), suggesting a much less efficient interaction. Furthermore, an interaction
231 between ChdC and IsdG, but not ChdC and IsdI, was also observed (Fig. 4E and 4F).
232 Together, these data reveal that IsdG interacts with the last two enzymes of the CPD pathway, namely
233 CpfC and ChdC.
234
235 The FLIM-FRET protein-protein interaction studies were complemented with an *in vitro* homogeneous
236 fluorescence polarization binding assay to monitor macromolecular complex formation between the
237 two recombinant *S. aureus* proteins, CpfC and IsdG. For this purpose, *S. aureus* CpfC was covalently
238 conjugated to dansyl chloride, a long-lived fluorescent probe, in order to evaluate its binding to IsdG
239 using both steady-state and time-resolved fluorescence anisotropy measurements (Valeur and
240 Berberan-Santos, 2012). The dansyl-labeled CpfC gave a high steady-state fluorescence anisotropy in
241 solution (<r> = 0.164 ± 0.002; Fig. 4G and Fig. S5). Upon increasing the concentration of dimeric IsdG in
242 solution, the steady-state fluorescence anisotropy of dansyl-CpfC steadily augmented (Fig. S6) reaching a
243 plateau level of <r> ~ 0.20 (Fig. 4G). In contrast, there was negligible binding observed between IsdI
244 and CpfC. Upon binding of dimeric IsdG to dansyl-labeled CpfC, the hydrodynamic volume of the protein
245 complex increased, slowing down the overall rotational tumbling of dansyl-labeled CpfC in solution
246 during its excited-state fluorescence lifetime and ultimately producing an increase in its steady-state
247 fluorescence anisotropy (Jameson and Ross, 2010; Valeur and Berberan-Santos, 2012). These data were
248 further corroborated by time-resolved polarized fluorescence measurements of dansyl-labeled CpfC
249 during its titration with IsdG. The fluorescence anisotropy decay of dansyl-labeled CpfC was greatly

affected by addition of IsdG to the solution (Fig. S6). In particular, the longer rotational correlation time, ϕ_2, assigned to the overall rotational motion of dansyl-labeled CpfC in solution, increased significantly from ϕ_2 ~ 67 ns for dansyl-labeled CpfC free in solution to ϕ_2 ~ 180 ns for dansy-labeled CpfC in the presence of 107 μM IsdG (Fig. 4H and Table S2). A binding K_d of 14.3 ± 2.6 μM (for the dimeric form of IsdG) and a $\langle r \rangle_B$ of 0.207 ± 0.003 was determined. These results clearly demonstrate that IsdG has the ability to bind to CpfC in solution at physiological pH.

Haemin increases IsdG expression

To assess the impact of haem uptake by the host to its endogenous haem biosynthesis, we determined the expression of the haem biosynthesis genes in *S. aureus* cells grown in medium supplemented with haemin (2.5 μM)**.** Addition of external haemin modified the mRNA content of *gtrR*, *uroD* and *chdC* (Fig. 5A). A similar change was observed with the genes encoding the iron-regulated surface determinants *isdA*, *isdB* and *isdI*, while *hrtB* and *htsA* were strongly upregulated (Fig. 5A). Treatment with exogenous haemin also caused a significant induction of the *isdCDEG* operon (Fig. 5A).

The pronounced change in *isdG* gene expression led us to analyse the amount of IsdG protein in cells containing excess haem, which was either added exogenously to the medium or produced endogenously through supplementation with aminolaevulinic acid. For this purpose, an IsdG-GFP fusion protein was constructed and the amount of IsdG was determined by the direct measurement of the fluorescence derived from the construct by flow cytometry. For comparative purposes, an IsdI-GFP fusion protein was also constructed. *S. aureus* expressing GFP, IsdG-GFP or IsdI-GFP were grown in either TSB supplemented with haemin or TSB containing an excess of aminolaevulinic acid. Addition of haemin increased the levels of IsdG, which was evidenced by the higher fluorescence of *S. aureus* cells expressing IsdG-GFP compared to cells expressing GFP alone (Fig. 5B). Cells expressing IsdI-GFP showed slightly higher fluorescence than cells expressing GFP alone, however at a much lower extent than cells containing IsdG-GFP (Fig. 5B). Interestingly, addition of aminolaevulinic acid caused an increase in fluorescence, revealing that internally produced haem also augments the cellular content of IsdG (Fig. 5C). In contrast, supplementation with aminolaevulinic acid produced a strong decrease of fluorescence in cells expressing IsdI-GFP. Therefore, the increase in the intracellular content of IsdG promoted by haem may trigger the interaction of IsdG with CpfC with a consequential impairment of the CPD pathway.

IsdG-like proteins are found in organisms that encode haem biosynthesis and lack haem uptake systems

IsdG-type haem degrading monooxygenases are a family of proteins that are characterized by the presence of an antibiotic biosynthesis monooxygenase (ABM) domain followed by a loop region

(Reniere *et al.*, 2011). To further test our proposal that IsdG-like proteins play a role in haem biosynthesis, we analysed the distribution of these proteins in complete and annotated genomes and their co-occurrence with genes encoding haem biosynthesis and/or haem uptake systems. For this purpose, an organism was considered to have a *de novo* haem biosynthesis ability if at least 75% of the pathway was encoded in its genome. For the PPD pathway, besides the 75% rule either CgdC or CgdH had to be present in the organism's genome. In addition, for organisms with the CPD pathway, ChdC had to be encoded in the genome. Also, an organism was considered to be able to perform haem uptake if its genome encoded a complete transport system with at least one haem binding protein (HmuT, IsdE or RV2003) (see Methods).

Data showed that among the so far complete and annotated genomes (5060) more than 1/5 of the organisms contain genes encoding IsdG-like proteins (Supplementary Data 1). The IsdG-like proteins are more represented in Gram-positive organisms (79%) (Fig. 6A). The majority of IsdG-like containing organisms encode complete (or almost complete) haem biosynthesis and uptake systems (Fig. 6B). Interestingly, IsdG-like proteins were not found in organisms that synthesize haem through the alternative haem biosynthesis pathway.

Within the 747 Gram-positive organisms that perform both haem biosynthesis and uptake, 422 (~56%) contained two or more genes coding for IsdG-like proteins. The remaining genomes contain only one IsdG-type protein but in ~50% of these organisms IsdG occurs together with genes coding for other haem monooxygenases belonging to the HemO, HmuO, ChuS/HmuS or HugZ families of haem monooxygenases (Fig. 6C). Within the 108 Gram-negative organisms, the majority (81%) encode two haem monooxygenases.

Noteworthy, IsdG-like proteins are found in organisms that lack haem uptake systems but contain haem biosynthesis pathways. Moreover, over 90% of the organisms that apparently have incomplete haem biosynthesis or uptake systems encode IsdG homologues together with haem biosynthesis ChdC and/or CpfC associated enzymes. These organisms belong to *Lactobacillus plantarum* or *Lactobacillus reitti* species that are unable to perform haem biosynthesis, but when grown with haem (and menaquinone), couple their haem-independent fermentative metabolism with aerobic respiration (Brooijmans *et al.*, 2009).

Altogether, the analysis showed that among the IsdG containing bacteria many of these organisms seem to rely only on haem biosynthesis pathways. Moreover, even the genomes that lack genes putatively involved in haem biosynthesis or uptake pathways still encode IsdG, ChdC and CpfC-like genes. Finally, the majority of the organisms with well-defined haem biosynthesis and uptake pathways contain two haem monooxygenases-like proteins, which supports the hypothesis that one haem monooxygenase could be used for haem degradation and the other could potentially interfere with haem biosynthesis, as shown in this work for *S. aureus*.

Discussion

We have investigated the way in which prokaryotes that obtain haem through haem biosynthesis and uptake pathways balance their high haem needs with the toxicity associated with free haem. We have shown that IsdG, one of the haem monooxygenases in the *S. aureus* haem uptake pathway, controls intracellular haem content through a protein-protein interaction with ferrochelatase that results in inhibition of iron-coproporphyrin chelatase activity. A comprehensive bioinformatics analysis showed that IsdG-like proteins are present in a significant number of prokaryotes that contain only the haem biosynthesis pathway. Altogether, the data allows us to propose that the IsdG protein family is not only the missing link between the two pathways, but also acts as the brake that avoids production of undesirably high intracellular haem levels.

In several microorganisms, IsdG-type haem monooxygenases, including the *S. aureus* IsdG and IsdI, have been assigned a haem-degradation role *in vitro* (Skaar *et al.*, 2004), and are required for growth when haemin is the only available iron source (Reniere and Skaar, 2008). We observed that IsdI impairs growth when haem is formed through PPD or CPD pathways, which is consistent with the haem degrading activity of IsdI. In contrast, IsdG inhibited growth only when haem was generated by the CPD pathway, which suggests that IsdG interferes with the haem synthesis pathway.
Experimentally, we have demonstrated that strains containing IsdG have lower amounts of intracellular haem formed via the haem biosynthesis route and IsdG has the ability to decrease the iron-coproporphyrin chelatase activity of *S. aureus* CpfC. Furthermore, we showed by *in vivo* FLIM-FRET and fluorescence polarization binding assays that the two proteins interact and have generated a model for this interaction.

Under our conditions, addition of haem did not repress expression of the *S. aureus* CPD pathway linked genes, as has been reported previously for several organisms. The only exception to this observation is with the *Corynebacterium diphtheriae gtrR*, which is repressed under haem replete conditions (Bibb *et al.*, 2007). Thus, in general, haem does not serve as a feedback factor to repress its own synthesis at the transcriptional level. Haem toxicity *in S. aureus* has previously been associated with induction of the

haem regulator transporter HrtAB (Torres *et al.*, 2007). In agreement with this, when we treated cells with haemin, they exhibited induction of the first gene of the *hrtBA* operon. More significantly, we observed that expression of *isdG* was highly upregulated, indicating that IsdG may play a role in haem regulation. Furthermore, this upregulation translated into an increase of the IsdG protein abundance, as shown by FACS. Consistent with our results, IsdG, but not IsdI, has been reported to be regulated at the post-transcriptional level such that it is stabilized in the presence of haem and undergoes proteolytic degradation in the absence of haem (Reniere and Skaar, 2008; Reniere *et al.*, 2011). We observed a higher abundance of IsdG in *S. aureus* when it was grown in iron-replete medium, while other authors have reported that IsdG is more expressed upon addition of exogenous haemin to iron-starved cultures of *S. aureus* (Reniere and Skaar, 2008). These differences are most probably due to the experimental conditions used in the two assays.

The bioinformatics analysis reported in this work estimates that among bacteria that only synthesize haem endogenously approximately one third contain IsdG-like proteins. Even in organisms with a less well-defined haem biosynthesis pathway, IsdG-type enzymes coexist with ChdC and/or CpfC enzymes (Fig. S6B). This is in agreement with previous studies that have shown that the presence of IsdG is not restricted to organisms that have both the haem biosynthesis and acquisition machineries (Sousa *et al.*, 2013). However, in many cases the genes encoding for IsdG-like enzymes are co-localised with genes for PPD and CPD pathways. Hence, IsdG may regulate both these routes for endogenous biosynthesis of haem. Our data support the proposal that in *S. aureus* this control occurs through the interaction of IsdG with CpfC, as the presence of IsdG strongly decreases ferrochelatase activity and the intracellular haem abundance. Although at this stage the interaction at the molecular level between IsdG and CpfC cannot be fully described due to the lack of detailed structural information, we hypothesize that this interaction may block the access of the substrate to the porphyrin binding cleft of CpfC.

Until now there has been a great deal of focus on the mechanisms of haem uptake as a source of iron, whilst the question of how microbes prevent unwanted haem toxicity by fine tuning exogenously acquired haem with haem synthesized endogenously remains less understood.

The scheme depicted in Fig. 7 summarises our current proposal for bacteria to maintain their intracellular free haem pool below hazardous levels. Organisms with both haem biosynthesis and uptake systems use external haem in two different ways. External haem is transported into the cytoplasm through dedicated systems, including those of the Isd-type, where it is either directly incorporated into apo-haemoproteins or degraded by IsdI, HemO, HmuO, ChuS/HmuS or HugZ- like proteins to release iron. On the other hand, haem also increases the abundance of IsdG to levels that are sufficient to interact with CpfC blocking its function, i.e, impairing the haem biosynthesis pathway. Hence, one of the haem monooxygenase enzymes would likely be dedicated to haem degradation to provide a source for

iron whilst the other would act to restrain internal haem biosynthesis. Moreover, in systems that only synthesize haem via an internal haem biosynthesis pathway and contain IsdG-like proteins, IsdG may play an important role in preventing excessive production of haem.

Collectively, our results show that the haem uptake and biosynthesis are not independent processes, and that IsdG-like proteins have a role in the crosstalk between these two systems which allow bacteria to adapt to a range of environments while avoiding haem toxicity. Therefore, we predict that the design of inhibitory drugs targeting the IsdG family of proteins will have a significant therapeutic benefit for the treatment of pathogenic infections.

Experimental Procedures

Strains and growth conditions

The strains used in this work are listed in Supplementary Table 3 and were grown under aerobic conditions, at 37 °C and 150 rpm. *S. aureus* was cultured in tryptic soy broth (TSB), with the exception of the cells that were used for RNA extraction that were cultured in Roswell Park Memorial Institute (RPMI) medium supplemented with 1% casamino acids. *E. coli* strains were grown in Luria-Bertani (LB), except *E. coli* expressing the pWhiteWalker plasmid that was grown in Brain Heart Infusion (BHI) medium. Selection was achieved by addition of the indicated antibiotics, namely kanamycin (50 μg ml^{-1}), erythromycin (20 μg ml^{-1} or 400 μg ml^{-1} for FRET experiments) and ampicillin (100 μg ml^{-1}).

RNA isolation and quantitative real-time RT-PCR assays

Overnight cultures of *S. aureus* were diluted to an optical density at 600 nm (OD_{600}) of 0.1 on RPMI containing 1% casamino acids or RPMI containing 1% casamino acids supplemented with haemn (2 μM). Cells at an OD_{600} = 1 were treated with an ice-cold ethanol/phenol RNA protective solution (5%), centrifuged at 2000 *x g* for 5 min, and the pellets flash frozen in liquid nitrogen. For RNA isolation, pellets were thawed on ice, resuspended in 10 mM Tris pH 8 and lysed with 2 mg ml^{-1} lysozyme and 30 μg ml^{-1} lysostaphin at 37 °C, for 30 min. The lysates were transferred to Aurum RNA Binding Mini Columns and total RNA was extracted using Aurum™ Total RNA Mini Kit (Bio-Rad), following the manufacturer´s instructions. Contaminating DNA was removed using Ambion® TURBO DNA-free™DNase kit (Life Technologies), and RNA concentration and purity were evaluated in a Nanodrop ND-1000 UV–visible spectrophotometer (Thermo Fisher Scientific).

For the cDNA synthesis, 800 ng of RNA was reverse transcribed with the Transcriptor High Fidelity cDNA Synthesis Kit (Roche) using the Anchored-oligo (dT)18 and Random Hexamer primers. Quantitative real-time RT-PCR assays were done in a LightCycler® 480 (Roche) using the oligonucleotides listed in Supplementary Table 4 and the LightCycler® 480 SYBR Green I Master kit (Roche). Relative quantification of each gene is shown in relation to the 16S rRNA reference gene, whose expression does not vary under the tested conditions, and using the comparative C_T method. Assays were done for two independent biological samples analysed in triplicate.

Gene cloning

The genes encoding *S. aureus* Newman IsdG and IsdI were amplified, by standard PCR reactions, from genomic DNA using the Phusion High-Fidelity DNA Polymerase (Thermo Fischer) and the oligonucleotides described in Supplementary Table 4. DNA fragments were cloned into either pET-23b or pET-28a vectors (Novagen) to produce wild type proteins and N-terminal His-Tag fused proteins, respectively. All plasmids were confirmed for gene integrity by DNA sequencing.

For the complementation experiments, plasmid pET-23b containing combinations of *cpfC*, *chdC*, *cgoX*, *isdG*, *isdI* of *S. aureus* (*Sa*), namely pET-23b-Sa*cgoXcpfCchdCisdG*, pET-23b-*SacgoXcpfCchdCisdI,* were generated by the link and lock methodology (McGoldrick *et al.*, 2005).

For the Förster Resonance Energy Transfer (FRET) studies, the *isdG*, *isdI*, *cpfC* and *chdC* genes amplified from *S. aureus* genomic DNA, as described above, were cloned into pBCB plasmids, which were kindly provided by Mariana Pinho (ITQB-NOVA, Portugal).

CpfC and ChdC proteins fused at the N-terminal to GFP were generated by cloning the respective genes into SphI/SpeI-pBCB1-*gfp* vector whereas IsdG, IsdI, CpfC and ChdC were fused at the C-terminal to mCherry (mCh) by cloning into the KpnI/NheI-pBCB7-*mCh* vector. After confirmation of the correct sequence of the fusion genes, BL21(DE3)Gold cells were used as recipient for the generated plasmids and analysed by FRET.

For the flow cytometry experiments, the *S. aureus isdG* gene was cloned into the EcoRI/KpnI restriction sites of pWhiteWalker3 (kind gift of Simon Foster, University of Sheffield, UK) to generate the in-frame fusion *isdG-gfp* (pWW-*isdG-gfp*). Plasmid pWhiteWalker3 that expresses only GFP (pWW-*gfp*) was also constructed to be used as control. For this purpose, a gBlock gene fragment of 244 bp that includes a ribosomal binding site and the N-terminal sequence of GFP (Integrated DNA Technologies) was cloned into EcoRI/NcoI-pWW-*isdG-gfp*. The correct sequence of the two recombinant plasmids was confirmed, and after electroporation into *S. aureus* the fluorescence level of GFP was analysed by flow cytometry.

Complementation experiments

Plasmid pET-23b harbouring combinations of genes of *S. aureus cgoX*, *cpfC*, *chdC*, *isdG*, *isdI,* and *E. coli ppfC* were transformed into competent cells of *E. coli* ferrochelatase mutant Δ*ppfC* (formerly Δ*visA*, kind gift from Mark O'Brian, State University of New York at Buffalo, New York, USA) (Frustaci and O'Brian, 1993). Overnight cultures were grown in LB supplemented with ampicillin and 5 µM haemin and 1 ml aliquots of cells were centrifuged at 1700 *x g* for 5 min and washed three times with LB. These pellets were used to inoculate LB-ampicillin, and growth was monitored for 8 h by measuring the optical density at 600 nm (OD_{600}) in a spectrophotometer (Multiskan™ GO, ThermoFisher Scientific). Assays were done for two independent biological samples.

Infection assays

Murine macrophages J774A.1 (5 x 10^5 cells/ml) (LGC Promochem) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% of fetal bovine serum, 100 µM of non-essential amino acids (Gibco), 50 U ml^{-1} of penicillin (Gibco), and 50 µg ml^{-1} of streptomycin (Gibco) in 24-well plates, at 37°C in a 5% CO2-air atmosphere. Prior to infection, macrophages were activated with 5 µg ml^{-1} LPS (Sigma) and 1 µg ml^{-1} IFN-Y (Sigma), for 5 h. *S. aureus* wild type and Δ*cpfC* mutant were inoculated, separately, in TSB medium and grown overnight. The cultures were re-inoculated in TSB, and 4 ml of cells grown to an OD_{600}= 0.4 were centrifuged at 4300 *x g* for 5 min, and washed three times with PBS. Cells were then resuspended in DMEM, diluted to OD=0.05 (~10^7 CFU ml^{-1}) and used to infect the murine macrophages. After 30 min of infection, at 37 °C, cells were washed twice with PBS and incubated with 50 µg ml^{-1} of gentamycin, at 37 °C, for 10 min, to prevent extracellular bacterial growth. Immediately after the addition of DMEM (time zero) and 2 and 4 h later, macrophages were collected, lysed with 2% of saponin and the number of intracellular bacteria was determined by CFU counting on TSB agar plates, which also contained 4 µM haemin to allow for the growth of the ferrochelatase mutant strain. Experiments were done for three independent biological samples assayed in duplicate.

Production of recombinant proteins

For the purification of *S. aureus* ferrochelatase CpfC, pET-23b-*SahemH* was transformed into BL21STAR(DE3) pLysS (Novagen) competent cells that were grown in LB medium at 37 °C and 150 rpm. Cells at an OD_{600} of 0.6 were induced by addition of 400 µM of isopropyl β-D-1-thiogalactopyranoside (IPTG) in the presence of 8 mg ml^{-1} of $FeSO_4$ and grown for an additional 16-20 h, at 20 °C. For the production of *S. aureus* IsdG and IsdI, plasmids pET-28b-*SaisdG* and pET-28b-*SaisdI* were transformed, separately, in competent cells of *E. coli* BL21(DE3)Gold. Cells were grown in LB medium until reaching an OD_{600}=0.7, and the protein expression was induced by addition of 1 mM of IPTG to the medium, and cells grown at 30°C for 3 h. Cells were harvested by centrifugation (10000 *x g*, 15 min, 4 °C), and the

pellets were resuspended in 50 mM Tris-HCl pH 7.5 (Buffer A). Cells were disrupted in a French press operating at 1000 Psi, and centrifuged at 27216 *x g*, at 4°C, for 15 min. The supernatant was applied onto a Ni-Sepharose chelating fast flow column (GE Healthcare), previously equilibrated with buffer A supplemented with 10 mM of imidazole, and the proteins were eluted at 400 mM of imidazole and dialysed against buffer A supplemented with NaCl (150-400 mM). Proteins with level of purity >95%, as judged by SDS-PAGE, were concentrated in an Amicon Stirred Ultrafiltration Cell using a 10 kDa membrane (Millipore) and frozen until use.

Enzymatic activities

Catalase activity

Overnight cultures of *S. aureus* wild type, and Δ*cpfC* mutants were grown in TSB only or supplemented with 4 µM haemin to an OD_{600} of 1. Cells were pelleted by centrifugation (10000 *x g*, 10 min), resuspended in buffer A and incubated with 7 µg of lysostaphin for 45 min, at 37 °C. The protein concentration of the cell lysates was determined using the Pierce Bicinchoninic acid Protein Assay (Thermo Scientific) and Sigma protein standards. For the catalase activity assays, approximately 26 µg of cells lysate proteins were added to buffer A containing 10 mM H_2O_2. The catalase activity was measured by following the decrease in absorbance at 240 nm for the consumption of H_2O_2 (ε_{240nm} = 43.6 M^{-1} cm^{-1}) in a Shimadzu UV-1700 spectrophotometer. Assays were done in triplicate.

Ferrochelatase activity

The assays were performed under anaerobic conditions in a Coy model A-2463 and Belle Technology chamber equipped with a Shimadzu UV-1800 spectrophotometer. For the preparation of coproporphyrin III (copro III) solution, 1–2 drops of 25% NH_4OH was added to 1-3 mg of copro III powder followed by addition of 1 ml of water. Copro III concentration was determined spectrophotometrically in 0.1 M HCl (ε_{548} = 16.8 mM^{-1} cm^{-1}). For the ferrochelatase assay, CpfC (0.3µM) was pre-incubated with IsdI or IsdG (1:10 molar ratio) at room temperature, for 10 min. Next, these proteins were added to the reaction mixture that contained copro III (10 µM) and $(NH4)_2Fe(SO_4)_2$ (10 µM) in buffer Tris-HCl pH 8.0. The IsdG titration experiments, were performed using IsdG concentrations of 0.3, 0.7, 1.7 and 3.4 µM, which correspond to 1:1, 1:2, 1:5 and 1:10 stoichiometry relative to the concentration of CpfC. The chelatase activity was measured by following the decrease in absorbance of copro III, measured at 392 nm (ε_{392nm} = 0.115 $\mu M^{-1}cm^{-1}$). Assays were done in triplicate.

Haem abundance assays

S. aureus wild type and Δ*isdG* mutant cells were grown for 8h and then diluted in TSB in the presence of 400 µM of aminolaevulinic acid. Cells were harvested by centrifugation (8000 x g, 5 min, 4 °C), washed and resuspended in buffer A, and incubated with lysostaphin, at 37 °C, for 45 min. The supernatant was collected by centrifugation and the intracellular haem was quantified essentially as previously described (Wolf *et al.*, 1984; Levicán *et al.*, 2007). Cell extracts (250 µL) were mixed with the same volume of a solution containing 0.5 M NaOH and 2.5 % Triton-X-100, and the haematin formation was determined measuring the absorbance at 575 nm in a spectrophotometer Shimadzu UV-1700. Haemin from Frontier Scientific was used as standard.

Mass spectrometry of tetrapyrrole products

The plasmids pET-23b containing combinations of *cgoX*, *cpfC*, *chdC*, *isdG* and *isdI* were inserted into *E. coli* ferrochelatase mutant Δ*ppfC*. Cells were prepared as described above for the complementation experiments, and inoculated in LB medium to an OD_{600}~0.05 and grown for 7 h. Cells were harvested by centrifugation at 10000 *x g* for 10 min, resuspended in buffer A, disrupted at 900 Psi in a French press, and centrifuged at 17000 *x g* for 30 min, at 4 °C. Protein content of the cell-free lysates was quantified using a Nanodrop ND-2000C (Thermo Scientific). Lysates with the equal protein concentration were treated for haem extraction. Briefly, proteins were precipitated by incubation of lysates with an acetone:HCl (19:1 vol/vol) mixture for 20 min, at room temperature, and removed by centrifugation at 14000 *x g* for 2 min. Following addition to supernatants of 1 ml of cold water, few milligrams of $(NH_4)_2SO_4$ (Panreac) and 300 µl of pure ethyl acetate (Sigma), haem was extracted from the organic phase after centrifugation at 14000 *x g* for 2 min. Samples were resolved by HPLC-MS on an Ace 5 AQ column attached to an Agilent 1100 series HPLC, equipped with a diode array detector and coupled to a micrOTOF-Q II (Bruker) mass spectrometer. Separation of the products was achieved by applying a gradient composed by 0.1% TFA and acetonitrile, at a flow rate of 0.2 ml min^{-1}. The column was first equilibrated with 20% solvent B, and after sample injection the concentration of solvent B was increased up to 100% during 50 min. Haem quantification was done by measuring the area of absorbance peak at 400 nm for m/z 616 (haem).

Flow cytometry

Overnight cultures of *S. aureus* RN4220 transformed with pWW-*isdG-gfp,* pWW-isdI-gfp or pWW-*gfp* were grown in TSB to an OD_{600}~1.5, diluted to an OD_{600} of 0.1 in TSB supplemented with 10 µM of IPTG, and grown for one extra hour. At this stage, cells were divided in 10 ml aliquots and grown for 4 h in the absence and in the presence of 5 µM haemin or 400 µM of aminolaevulinic acid. Cells (1 ml) were

collected by centrifugation at 11400 *x g* for 1 min, and the pellets were washed 3 times with PBS, diluted in PBS to an OD_{600} of 0.1 and analysed in a Cell Sorter S3e™ (Biorad). For each sample, at least, 300,000 cells were collected and analysed with the FlowJo programme (Tree Star), and three biological samples were measured for each condition.

Fluorescent labelling of CpfC

S. aureus CpfC was covalently labelled with 5-dimethylaminonaphthalene-1-sulfonyl chloride (dansyl chloride) according to the manufacturer´s instructions (Invitrogen). Briefly, recombinantly produced and purified *S. aureus* CpfC (7.5 mg ml^{-1}), solubilized in 0.1 M sodium bicarbonate pH 8.6, was incubated in a 1:1 ratio with dansyl chloride (10 mg ml^{-1} in dimethylformamide) for 1 h at 4°C, in the dark and under continuous stirring. The reaction was quenched by the addition of hydroxylamine (1.5 M, pH 8.5), and the mixture was then loaded onto a PD-10 column (GE Healthcare) equilibrated with buffer A in order to remove the excess of free dye by gel filtration. The CpfC:dansyl labelling ratio was determined by spectrophotometric quantification of the dye (ε_{331nm} = 4,000 M^{-1} cm^{-1}) (Gustiananda *et al.*, 2004) and of CpfC (ε_{280nm} = 47,700 M^{-1} cm^{-1}) (Gasteiger *et al.*, 2005), and estimated to be 0.97.

Pull down assays

Interaction of CpFc and Isds proteins was investigated by pull-down assays. For this purpose, recombinant N-terminal polyhistidine tagged-CpfC was expressed from pET-23b plasmid and purified as described in the previous section. In order to immobilize the overproduced His-tagged-CpfC, 14 mg of protein was loaded onto a Ni-Sepharose fast flow column (GE Heathcare). The column was next washed with 10 mM Tris–HCl, pH 8.0 containing several concentrations of imidazole, and the His-tagged CpfC was eluted at 60 mM imidazole.

In parallel, *E. coli* BL21(DE3)Gold supernatants expressing non-His tagged versions of *S. aureus* IsdG and IsdI (pET-23b) were prepared, separately, as described above. The supernatants were loaded into the CpfC-bound Ni-Sepharose chelating column and washed with buffer 10 mM Tris–HCl, pH 8.0, and proteins that were subsequently eluted with 60 mM imidazole were analysed by SDS-PAGE. To discard non-specific interactions, control experiments were done similarly using the same supernatants loaded into columns with no bounded CpfC-His protein.

Steady-state fluorescence anisotropy measurements

Dansyl-labeled CpfC (0.67 µM) was incubated with variable concentrations of IsdG or IsdI in buffer A.

The steady-state fluorescence anisotropy of each sample, <*r*>, was calculated according to:

$$< r >= \frac{I_{\mathrm{VV}} - G \cdot I_{\mathrm{VH}}}{I_{\mathrm{VV}} + 2\,G \cdot I_{\mathrm{VH}}} \quad (1$$

where I_{VV} and I_{VH} are the fluorescence intensities (blank subtracted) of the vertically and horizontally polarized emission, when the sample is excited with vertically polarized light, respectively. The *G* factor ($G = I_{HV}/I_{HH}$) is an instrument correction factor which takes into account the transmission efficiency of the monochromator to the polarization of the light. Measurements were performed at 25 ^{o}C on a Fluorolog-3-21 spectrofluorometer (Horiba Jobin Yvon) with automated dual polarizers using 5-mm path length quartz cuvettes. The excitation wavelength was 340 nm with a bandwidth of 6 nm, and the fluorescence emission was recorded at 530 nm with 5 nm bandwidth. *Kd* and $<r>_B$ parameters were obtained by fitting the experimental data (steady-state fluorescence anisotropy, <*r*> versus $[\mathrm{P2}]_t$, the total concentration of the binding partner 2 (dimeric IsdG or IsdI)) using the equation:

$$< r >=< r >_f + \frac{([\mathrm{P1}]_t+[\mathrm{P2}]_t+K_\mathrm{d}) - \sqrt{([\mathrm{P1}]_t+[\mathrm{P2}]_t+K_\mathrm{d})^2 - 4[\mathrm{P1}]_t \cdot [\mathrm{P2}]_t} \cdot (<r>_b - <r>_f)}{2[\mathrm{P2}]_t}$$

where $[P1]_t$ represents the total concentration of the binding partner 1. The concentration of CpfC was fixed to 0.67 µM during the non-linear regression and assumed that the fluorescently-labelled protein is monomeric in solution; $[P2]_t$ is the total concentration of the binding partner 2. Binding stoichiometry was considered 1:1; $< r >_f$ represents the steady-state fluorescence anisotropy of the free protein. $<r>_F$= 0.165 (fixed during the analysis); and $< r >_b$ the steady-state fluorescence anisotropy of the bound protein.

Förster Resonance Energy Transfer (FRET) and Fluorescence Lifetime Imaging Microscopy (FLIM)

E. coli BL21(DE3)Gold were co-transformed with plasmids that express combinations of CpfC, ChdC, IsdI and IsdG fused to GFP and mCherry (mCh) fluorophore proteins, namely pBCB-*cpfC-gfp* and pBCB-*isdG-mCh*; pBCB-*cpfC-gfp* and pBCB-*isdI-mCh*; pBCB-*chdC-gfp* and pBCB-*isdG-mCh*; and pBCB-*chdC-gfp* and pBCB-*isdI-mCh* (Supplementary Table 1). Cells grown overnight in LB were sub-cultured into LB medium to an OD_{600} of 0.15, supplemented with 1 mM IPTG and grown for 4 h. Cells (2 ml) were centrifuged at 4300 *x g* for 3 min, and fixed by incubation with 4% formaldehyde (vol/vol) in PBS, at room temperature, for 30 min and 90 rpm, washed with PBS and immobilized on 8 well µ-slides (Ibidi, slides, Germany) for FRET-FLIM experiments.

All measurements were acquired in a Leica TCS SP5 (Leica Microsystems CMS GmbH, Mannheim, Germany) inverted confocal microscope (DMI6000). A 63x apochromatic water immersion objective with a NA of 1.2 (Zeiss, Jena Germany) was used for all experiments. GFP and mCherry were excited respectively with the 476 nm and 514 nm lines from an Argon laser, and fluorescence emission was collected in the 485-540 nm range for GFP, and 580-700 nm for mCherry. In these conditions, spectral bleed-through was negligible.

Fluorescence lifetime imaging microscopy (FLIM) measurements were performed by time correlated single photon counting (TCSPC) using the confocal microscope coupled to a multiphoton Titanium: Sapphire laser (Spectra-Physics Mai Tai BB, Darmstadt, Germany) as the excitation source. FLIM data was acquired during 90-180 seconds to achieve reasonable photon statistics. The excitation wavelength was set to 840 nm and emission light was selected with a dichroic beam splitter with an excitation SP700 short-pass filter and an emission 525±25 nm band-pass filter inserted in front of the photomultiplier. Images were acquired using a Becker and Hickl SPC 830 module. Fluorescence decays for each cell were calculated by integrating the FLIM data for all pixels of each individual cell. Fluorescence lifetimes were obtained by analysing the fluorescence decays through a least square iterative re-convolution of decay functions with the instrument response function (IRF) using the software SPCImage (Becker and Hickl, Berlin, Germany). Intensity-weighted mean fluorescence lifetime ($\langle\tau\rangle$) of multiexponential decays were calculated as $\langle\tau\rangle = \sum_i \alpha_i \tau_i$, were α_i are the pre-exponential factors and τ_i are the individual lifetime values. Average FRET efficiencies in each cell can be determined from $\langle E\rangle = 1 - \langle\tau\rangle_{DA}/\langle\tau\rangle_D$, where $\langle\tau\rangle_{DA}$ and $\langle\tau\rangle_D$ are the donor intensity-weighted mean fluorescence lifetime in the presence or absence of acceptor, respectively. Assays were done for two independent biological samples and FRET efficiencies were determined for several cells.

Genomic analysis

A dataset composed of 5060 complete prokaryotic genomes was downloaded from RefSeq (June 2016) (Pruitt *et al.*, 2004). These correspond to all complete genomes available at the time in RefSeq. Taxonomic information was retrieved from NCBI and genomes were grouped by phylum or class. Homologous proteins involved in the several steps of haem biosynthesis and uptake (Urogen III synthesis: GtrR, GsaM, PbgS, HmbS and UroS; PPD haem biosynthesis UroD-(CgdC/CgdH)-(PgoX/PgdH1/PgdH2)-PpfC; alternative haem biosynthesis: AhbABCD; *S. aureus* haem biosynthesis variant: UroD-CgoX-CpfC-ChdC; *Staphylococcus* haem uptake and degradation: HstABC, IsdABCDEFGHI-SrtB, *B. subtilis* haem uptake: IsdCEDFGX1X2-Hal; haem monooxygenases IsdG-type: MhuD, HmuQ/D; other haem monooxygenases: HemO, HmuO, HmuS, HugZ, ChuS; haem transporters: HmuTUV, IsdEFD, HtsABC, Rv2002c-Rv2003-Rv2006c) were identified by BLAST (Altschul *et al.*, 1997) (E-value smaller than 10^{-10} and local amino acid identity of at least 25%). To distinguish between haem transporters systems

from siderophores and/or cobalamin importers, sequences from biochemically-characterized transporters were used as queries as well. All query proteins are listed in Supplementary Data 1. When relevant, PFAM-A domain annotations were obtained by using the HMM approach as available at PFAM (Finn *et al.*, 2016). Query coverage, gene fusions and genomic organization were also used for the identification of true positive hits. Due to their small length, to distinguish between the different IsdG homologous, an all versus all blast of putative IsdGs and query sequences and an alignment were performed. Hits were classified according to their best hit based on sequence identity, conservation of the catalytic triad (for HmoA) and sequence length. Selected IsdG-type homologous identified by Blast were aligned with ClustalO (Sievers *et al.*, 2011) and a maximum likelihood tree reconstructed with IQTree (Best-model selection WAG+G4)(Nguyen *et al.*, 2015).

To distinguish between "bona fide" CgdH from highly similar non-functional CgdHs, all genomes were also queried for the presence of non-functional CgdH in a similar way as in Dailey *et al* 2015 (Dailey *et al.*, 2015). To distinguish between ChdC and chlorite dismutases, sequences were aligned with ClustalO and a phylogenetic tree performed. Sequences corresponding to chlorite dismutases were discarded based on the conservation of the catalytic residues and position within a phylogenetic tree. A haem biosynthesis pathway was only considered to be present in a genome if: 1) 80% of its genes were identified and 2) if the characteristic proteins were present, CgdH and/or CgdC for classical haem pathway and ChdC for *S. aureus* variant. Cases in which a variant of the canonical pathways above described were present and/or missing one gene were assigned as hybrid or incomplete. Due to the high sequence similarity between genes involved in haem d1 biosynthesis with genes from the haem alternative pathway (Bali *et al.*, 2011), in organisms containing cd_1 nitrite reductase, the haem alternative pathway identified genes were considered to be involved in haem d_1 synthesis instead.

Statistical analysis

In all figures, error bars represent the standard deviation of at least two biological samples. Statistical differences were calculated by the two-tailed Student's t-test using GraphPad Prism (GraphPad Software).

Acknowledgements

We are grateful to Mark O´Brian (University of Buffalo), Mariana Pinho (Instituto Tecnologia Química e Biológica António Xavier), and Simon Foster (University of Sheffield) for their generous gift of the Δ*visA*

strain, pBCB and pWhiteWalker plamids, respectively. We also thank Patrícia Ferreira for technical support. This work was financially supported by: Project LISBOA-01-0145-FEDER-007660 (Microbiologia Molecular, Estrutural e Celular) funded by FEDER funds through COMPETE2020 - Programa Operacional Competitividade e Internacionalização (POCI) and by national funds through FCT - Fundação para a Ciência e a Tecnologia" for grants PTDC/BBB-BQB/5069/2014, SFRH/BD/95912/2013 (MAMV) and SFRH/BD/118545/2016 (LSOS).

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389–402.

Anzaldi, L.L., and Skaar, E.P. (2010) Overcoming the heme paradox: heme toxicity and tolerance in bacterial pathogens. *Infect Immun* **78**: 4977–4989.

Bali, S., Lawrence, A.D., Lobo, S.A., Saraiva, L.M., Golding, B.T., Palmer, D.J., *et al.* (2011) Molecular hijacking of siroheme for the synthesis of heme and *d1* heme. *Proc Natl Acad Sci U S A* **108**: 18260–5.

Bibb, L.A., Kunkle, C.A., and Schmitt, M.P. (2007) The ChrA-ChrS and HrrA-HrrS signal transduction systems are required for activation of the *hmuO* promoter and repression of the *hemA* promoter in *Corynebacterium diphtheriae*. *Infect Immun* **75**: 2421–2431.

Brooijmans, R., Smit, B., Santos, F., Riel, J. van, Vos, W.M. de, and Hugenholtz, J. (2009) Heme and menaquinone induced electron transport in lactic acid bacteria. *Microb Cell Fact* **8**: 28.

Choby, J.E., and Skaar, E.P. (2016) Heme synthesis and acquisition in bacterial pathogens. *J Mol Biol* 16–18.

Dailey, H.A., Dailey, T.A., Gerdes, S., Jahn, D., Jahn, M., O'Brian, M.R., and Warren, M.J. (2017) Prokaryotic heme biosynthesis: multiple pathways to a common essential product. *Microbiol Mol Biol Rev* **81**: e00048-16.

Dailey, H.A., Gerdes, S., Dailey, T.A., Burch, J.S., and Phillips, J.D. (2015) Noncanonical coproporphyrin-dependent bacterial heme biosynthesis pathway that does not use protoporphyrin. *Proc Natl Acad Sci U S A* **112**: 2210–5.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., *et al.* (2016) The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res* **44**: D279-85.

Frustaci, J.M., and O'Brian, M.R. (1993) The *Escherichia coli visA* gene encodes ferrochelatase, the final enzyme of the heme biosynthetic pathway. *J Bacteriol* **175**: 2154–2156.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*. pp. 571–607.

Gustiananda, M., Liggins, J.R., Cummins, P.L., and Gready, J.E. (2004) Conformation of prion protein repeat peptides probed by FRET measurements and molecular dynamics simulations. *Biophys J* **86**: 2467–2483.

Hammer, N.D., Cassat, J.E., Noto, M.J., Lojek, L.J., Chadha, A.D., Schmitz, J.E., *et al.* (2014) Inter- and intraspecies metabolite exchange promotes virulence of antibiotic-resistant *Staphylococcus aureus*. *Cell Host Microbe* **16**: 531–537.

Hammer, N.D., Reniere, M.L., Cassat, J.E., Zhang, Y., Hirsch, A.O., Hood, M.I., and Skaar, E.P. (2013) Two heme-dependent terminal oxidases power *Staphylococcus aureus* organ-specific colonization of the vertebrate host. *MBio* **4**: 1–9.

Heinemann, I.U., Jahn, M., and Jahn, D. (2008) The biochemistry of heme biosynthesis. *Arch Biochem*

Biophys **474**: 238–251.

Jameson, D.M., and Ross, J.A. (2010) Fluorescence Polarization/Anisotropy in Diagnostics and Imaging. *Chem Rev* **110**: 2685–2708.

Levicán, G., Katz, A., Armas, M. de, Núñez, H., and Orellana, O. (2007) Regulation of a glutamyl-tRNA synthetase by the heme status. *Proc Natl Acad Sci U S A* **104**: 3135–3140.

Lobo, S.A.L., Scott, A., Videira, M.A.M., Winpenny, D., Gardner, M., Palmer, M.J., *et al.* (2015) *Staphylococcus aureus* haem biosynthesis: characterisation of the enzymes involved in final steps of the pathway. *Mol Microbiol* **97**: 472–487.

Lobo, S.A.L., Warren, M.J., and Saraiva, L.M. (2012) Sulfate-reducing bacteria reveal a new branch of tetrapyrrole metabolism. *Adv Microb Physiol* **61**: 267–295.

Mayfield, J. a., Hammer, N.D., Kurker, R.C., Chen, T.K., Ojha, S., Skaar, E.P., and DuBois, J.L. (2013) The chlorite dismutase (HemQ) from *Staphylococcus aureus* has a redox-sensitive heme and is associated with the small colony variant phenotype. *J Biol Chem* **288**: 23488–23504.

Mazmanian, S.K., Skaar, E.P., Gaspar, A.H., Humayun, M., Gornicki, P., Jelenska, J., *et al.* (2003) Passage of heme-iron across the envelope of *Staphylococcus aureus*. *Science* **299**: 906–9.

McGoldrick, H.M., Roessner, C.A., Raux, E., Lawrence, A.D., McLean, K.J., Munro, A.W., *et al.* (2005) Identification and characterization of a novel vitamin B12 (cobalamin) biosynthetic enzyme (CobZ) from *Rhodobacter capsulatus*, containing flavin, heme, and Fe-S cofactors. *J Biol Chem* **280**: 1086–1094.

McNicholas, P.M., Javor, G., Darie, S., and Gunsalus, R.P. (1997) Expression of the heme biosynthetic pathway genes *hemCD*, *hemH*, *hemM* and *hemA* of *Escherichia coli*. *FEMS Microbiol Lett* **146**: 143–148.

Nguyen, L.T., Schmidt, H.A., Haeseler, A. Von, and Minh, B.Q. (2015) IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* **32**: 268–274. Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2004) NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res* **33**: D501–D504. Reniere, M.L., Haley, K.P., and Skaar, E.P. (2011) The flexible loop of *Staphylococcus aureus* IsdG is required for its degradation in the absence of heme. *Biochemistry* **50**: 6730–7.

Reniere, M.L., and Skaar, E.P. (2008) *Staphylococcus aureus* haem oxygenases are differentially regulated by iron and haem. *Mol Microbiol* **69**: 1304–1315.

Reniere, M.L., Ukpabi, G.N., Harry, S.R., Stec, D.F., Krull, R., Wright, D.W., *et al.* (2010) The IsdG-family of haem oxygenases degrades haem to a novel chromophore. *Mol Microbiol* **75**: 1529–1538.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.

Skaar, E.P., Gaspar, A.H., and Schneewind, O. (2004) IsdG and IsdI, heme-degrading enzymes in the cytoplasm of *Staphylococcus aureus*. *J Biol Chem* **279**: 436–443.

Skaar, E.P., and Schneewind, O. (2004) Iron-regulated surface determinants (Isd) of *Staphylococcus aureus*: Stealing iron from heme. *Microbes Infect* **6**: 390–397.

Skaar EP, Humayun M, Bae T, DeBord KL, S.O. (2004) Iron-source preference of *Staphylococcus aureus* infections. *Science* **305**: 1626–8.

Sousa, F.L., Thiergart, T., Landan, G., Nelson-Sathi, S., Pereira, I.A.C., Allen, J.F., *et al.* (2013) Early bioenergetic evolution. *Philos Trans R Soc B Biol Sci* **368**: 20130088–20130088.

Stauff, D.L., Bagaley, D., Torres, V.J., Joyce, R., Anderson, K.L., Kuechenmeister, L., *et al.* (2008) *Staphylococcus aureus* HrtA Is an ATPase required for protection against heme toxicity and prevention of a transcriptional heme stress response. *J Bacteriol* **190**: 3588–3596.

Tarai, B., Das, P., and Kumar, D. (2013) Recurrent challenges for clinicians: emergence of methicillin-resistant *Staphylococcus aureus*, vancomycin resistance, and current treatment options. *J Lab Physicians* **5**: 71.

Thöny-Meyer, L. (1997) Biogenesis of respiratory cytochromes in bacteria. *Microbiol Mol Biol Rev* **61**: 337–76.

Torres, V.J., Stauff, D.L., Pishchany, G., Bezbradica, J.S., Gordy, L.E., Iturregui, J., *et al.* (2007) A *Staphylococcus aureus* regulatory system that responds to host heme and modulates virulence. *Cell Host Microbe* **1**: 109–19.

Valeur, B., and Berberan-Santos, M.N. (2012) *Molecular Fluorescence: Principles and Applications, Second Edition*. .

Wang, L.Y., Brown, L., Elliott, M., and Elliott, T. (1997) Regulation of heme biosynthesis in *Salmonella typhimurium*: Activity of glutamyl-tRNA reductase (HemA) is greatly elevated during heme limitation by a mechanism which increases abundance of the protein. *J Bacteriol* **179**: 2907–2914.

Wolf, H., Lang, W., and Zander, R. (1984) Alkaline haematin D-575, a new tool for the determination of haemoglobin as an alternative to the cyanhaemiglobin method. I. Description of the method. *Clin Chim Acta* **136**: 95–104.

813

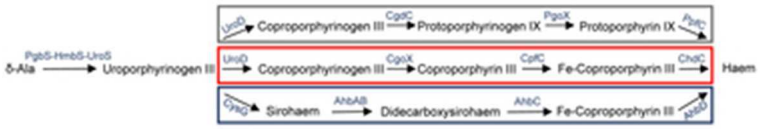


Figure 1. Three different haem biosynthesis pathways that diverge from uroporphyrin III. The enzymes in the three pathways are conserved between aminolaevulinic acid (δ-Ala) and uroporphyrinogen III. The protoporphyrin dependent (or classical) pathway is boxed in grey and converts uroporphyrinogen III into haem using protoporphyrin as an intermediate. The CPD pathway is represented in red and the sirohaem dependent pathway (alternative haem biosynthesis pathway) highlighted in blue.

32x5mm (300 x 300 DPI)

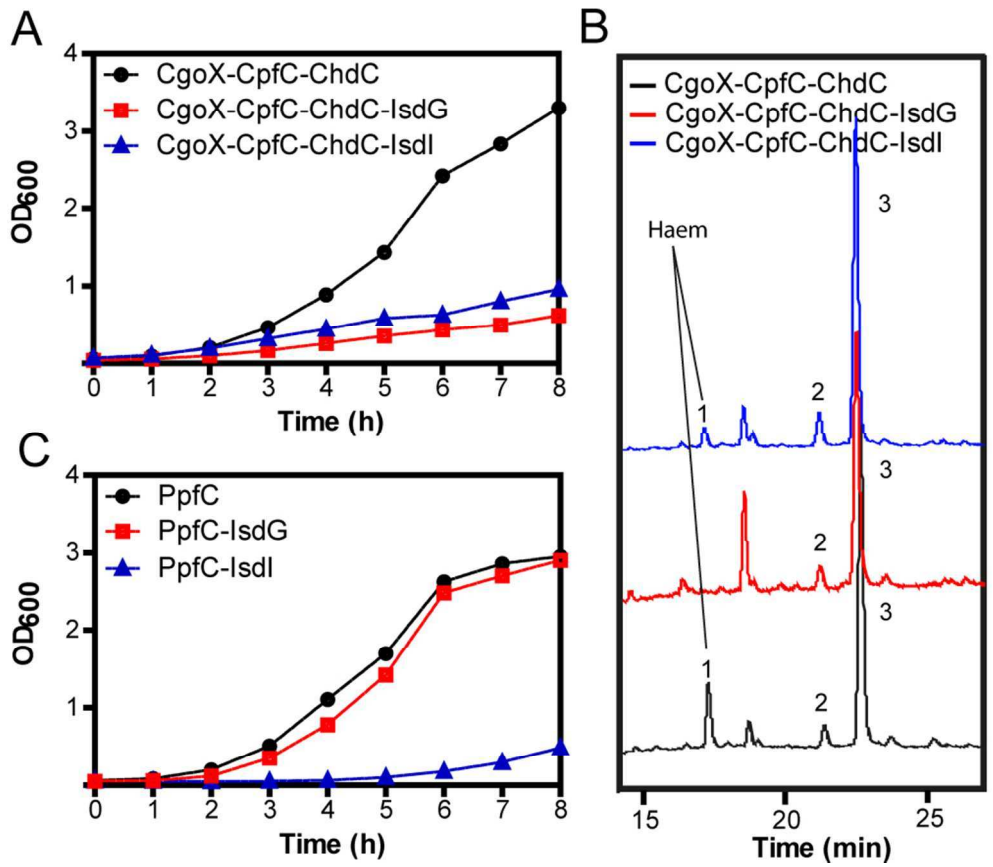


Figure 2. S. aureus IsdG and IsdI impair haem formation. (A) Growth of E. coli ΔppfC cells complemented with S. aureus CgoX-CpfC-ChdC alone, in the presence of S. aureus IsdG (CgoX-CpfC-ChdC-IsdG) or S. aureus IsdI (CgoX-CpfC-ChdC-IsdI). (B) HPLC-MS profile of cells depicted in (A), with peaks corresponding to haem (m/z = 616; peak 1), iron-coproporphyrin III (m/z = 708; peak 2), and protoporphyrin IX (m/z = 553; peak 3). (C) Growth of E. coli ΔppfC cells complemented with E. coli PpfC alone and together with S. aureus IsdG (PpfC-IsdG) or S. aureus IsdI (PpfC-IsdI). Experiments were performed for two independent biological samples.

86x74mm (300 x 300 DPI)

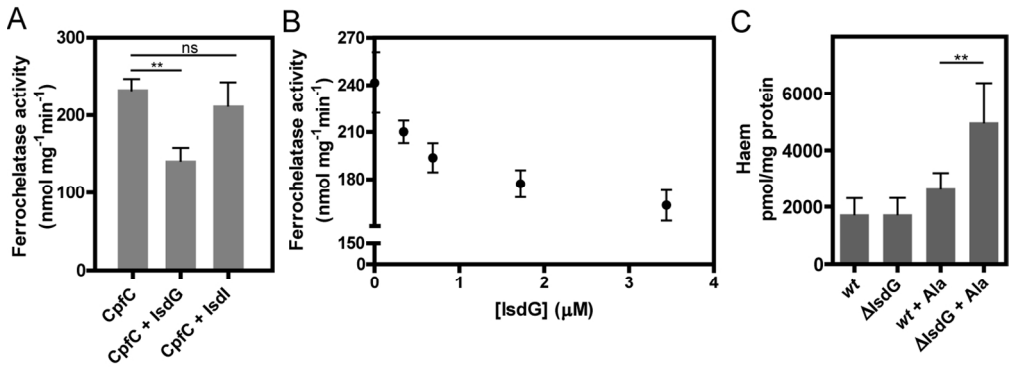


Figure 3. Influence of S. aureus IsdG on the CPD pathway. (A) IsdG decreases the ferrochelatase activity of S. aureus CpfC while IsdI does not impair the activity. (B) Inhibition of S. aureus CpfC ferrochelatase activity increases with IsdG concentration. (C) Intracellular haem quantification of S. aureus wild type and ΔisdG mutant for cells grown in the presence of 400 µM of aminolaevulinic acid (Ala).
In (A) and (B), activities were measured in reaction mixtures containing CpfC alone, or with the addition of IsdG or IsdI, using coproporphyrin III (10 µM) and (NH4)2Fe(SO4)2 (10 µM) as substrates. In (B), IsdG was used in the following concentrations: 0.3, 0.7, 1.7 and 3.4 µM. Data depict the mean and standard deviation of three samples using a two-tailed unpaired Student's t-test (** $p < 0.01$).

123x45mm (300 x 300 DPI)

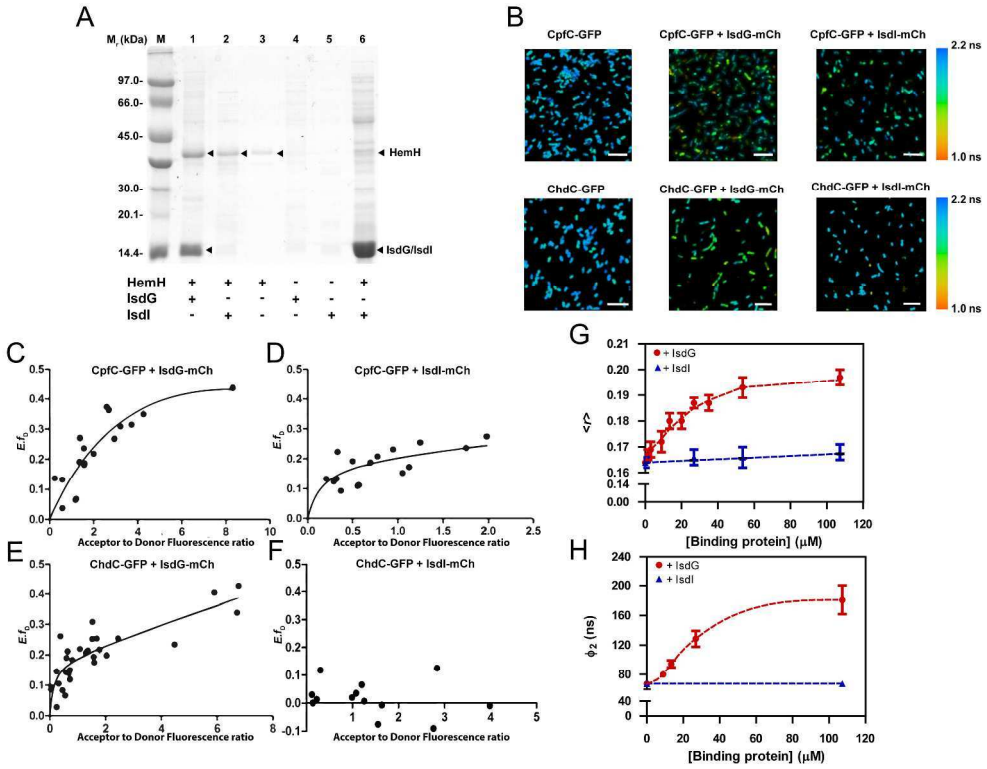


Figure 4. S. aureus CpfC and IsdG interact in vivo. (A) Ni-Sepharose pull down assays using CpfC-His tagged protein. Supernatants of E. coli expressing, separately, IsdG and IsdI proteins were loaded into CpfC-His bound or unbound columns and pulled down with imidazole. Figure depicts the protein molecular mass marker (M), protein fractions eluted at 60 mM imidazole (lanes 1 to 5), and protein fraction eluted with 10 mM Tris–HCl pH 8.0 following load with IsdI expressing supernatant into a CpfC-His tag bound column (lane 6). (B) Representative FLIM-FRET data of E. coli cells co-expressing CpfC-GFP or ChdC-GFP with IsdG-mCherry or IsdI-mCherry, shown as false colour lifetime images. Average fluorescence lifetime (<τ>) was rendered as colour according to the colour index indicated on the right. Scale bar represents 10 μm. For each cell in (B), the calculated average FRET efficiency (E) is plotted against the ratio of fluorescence (as measured by confocal microscopy) from the acceptor (mCherry) construct over donor (GFP), for CpfC-GFP/IsdG-mCherry (C), CpfC-GFP/IsdI-mCherry (D), ChdC-GFP/IsdG-mCherry (E), and ChdC-GFP/IsdI-mCherry (F). Lines are drawn as a guide to the eye. Data are from one representative sample. (G) Steady-state and (H) time-resolved fluorescence polarization binding assay between dansylated CpfC and IsdG or IsdI. (G) The steady-state fluorescence anisotropy, <r>, and (H) long rotational correlation time, ϕ2, of dansylated CpfC are plotted as a function of IsdG (red circles) and IsdI (blue triangles) concentrations expressed as monomers. Conditions consisted of 0.7 μM dansyl-labeled CpfC in 50 mM Tris-HCl buffer pH 8.0, at 25 °C. Dansyl fluorescence was monitored at 530 nm with excitation at 340 nm. The cuvette path length was 5 mm. Data represent the mean values of three independent measurements and error bars represent the standard deviation.

273x210mm (300 x 300 DPI)

A

Gene	Fold Change
gtrR	2.7 ± 0.5
uroD	2.6 ± 0.8
chdC	2.9 ± 0.1
isdA	2.4 ± 0.8
isdB	2.5 ± 0.2
isdE	6.5 ± 1.4
isdC	4.8 ± 0.4
hrtB	27.9 ± 0.5
htsA	48.5 ± 2.5
isdG	5.2 ± 1.0
isdI	2.2 ± 0.3

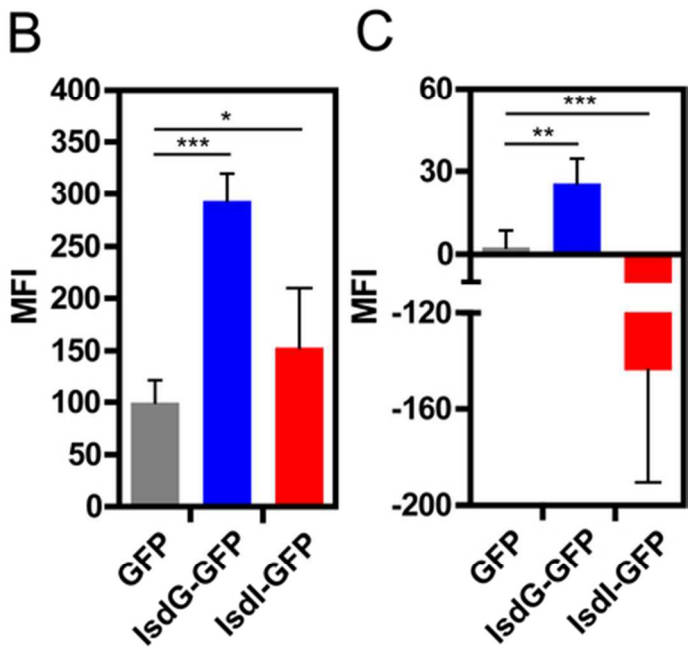


Figure 5. The cellular content of IsdG in S. aureus increases upon incubation with haemin. (A) Fold change in abundance of genes coding for enzymes involved in the haem biosynthesis and uptake pathways upon exposure to haemin. (B) Difference of the median GFP fluorescence intensity (MFI) between haemin-treated (5 μM) and untreated cells. (C) Difference of the median GFP fluorescence intensity between aminolaevulinic acid -treated (400 μM) and untreated cells. Fluorescence was measured in a flow cytometer and over 300,000 cells were counted. Data represent the mean and standard deviation of four measurements, using the two-tailed unpaired Student's t-test (*** $p < 0.001$, **$p=0.005$ *$p=0.04$).

56x36mm (300 x 300 DPI)

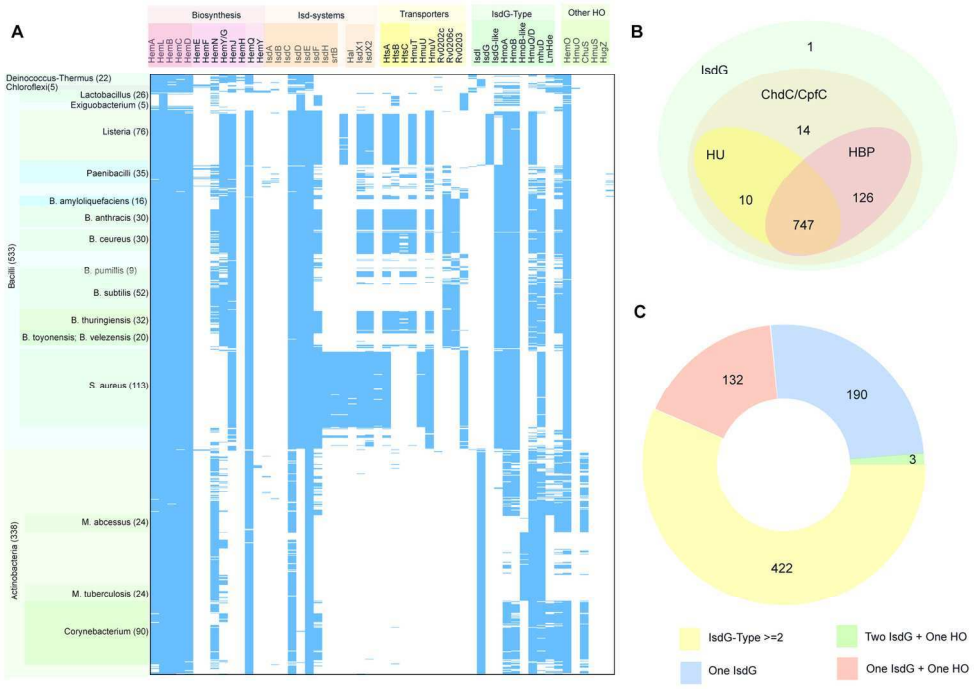


Figure 6. Distribution of haem metabolism and related genes across Gram-positive organisms containing IsdG-type enzymes. (A) Heat map representing the presence (blue ticks) or absence (white) of genes (columns) involved in haem biosynthesis, Isd systems, haem uptake and degradation in the 898 genomes (rows) containing IsdG-type enzymes. The full list is given in Supplementary Data 1. (B) Co-occurrence of IsdG-like enzymes (green) with haem biosynthesis pathway HBP (red), haem uptake system (yellow) and ChdC/CpfC (brown). (C) Co-occurrence of additional haem monooxygenases in the 747 Gram-positive organisms able to uptake and de novo synthesise haem and that contain one or more than one IsdG-type enzymes.

148x108mm (300 x 300 DPI)

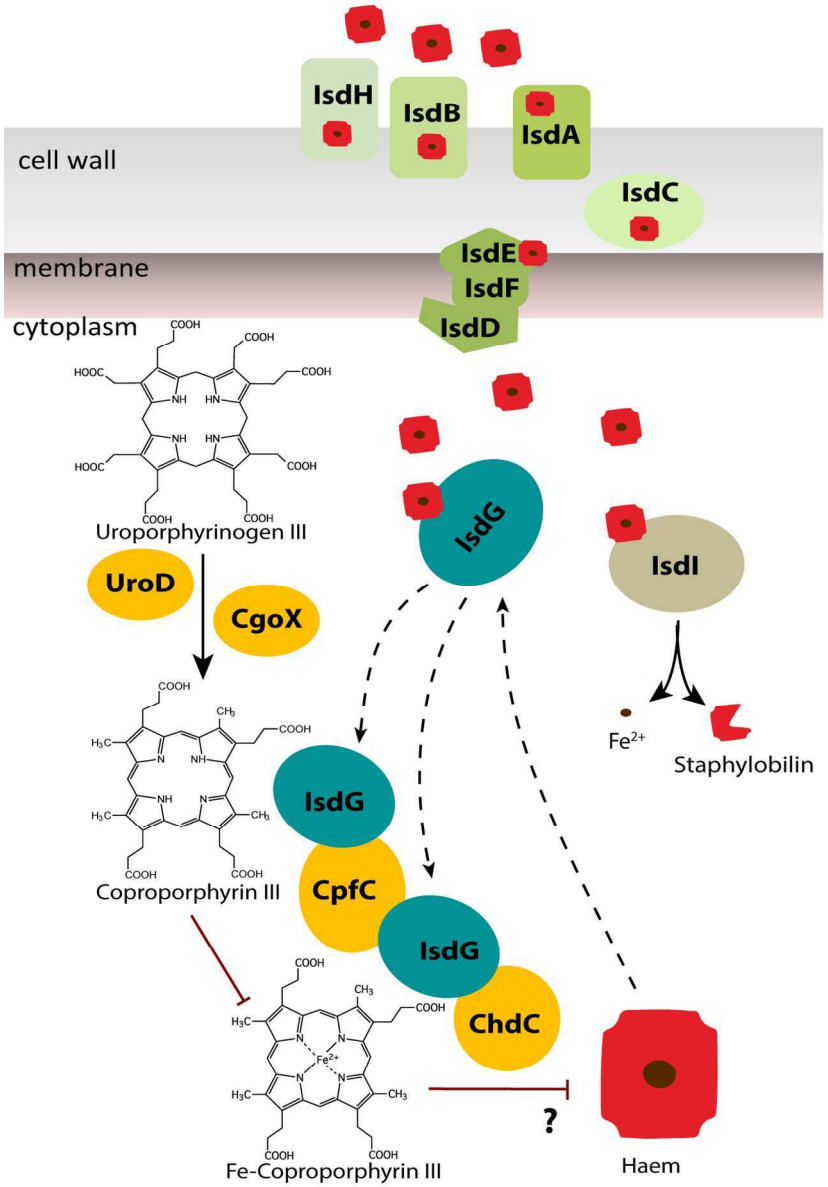


Figure 7. Proposed scheme for the role of IsdG-like proteins in the control of haem production through crosstalk between haem biosynthesis and uptake systems. S. aureus CPD is represented proceeding from uroporphyrinogen III to haem via the UroD, CgoX, CpfC and ChdC enzymes. S. aureus haem uptake is represented by the Isd system. In this scheme, IsdH and IsdB scavenge haemoglobin haem which enters into the cell through IsdA, IsdC and IsdEFD. In the cytoplasm, haem is mainly degraded by IsdI to staphylobilin, iron and formaldehyde. Increase of haem content augments the IsdG abundance to levels that allow its interaction with CpfC, resulting in the inhibition of the CPD pathway.

111x163mm (300 x 300 DPI)